Conditionarea rădăcinilor ecuaţiilor polinomiale

Radu Trîmbiţaş

30 aprilie 2020

1 Senzitivitatea rădăcinilor polinomiale

Vom considera condiţionarea unei rădăcini ξ a unei ecuaţii polinomiale $p(x)$. Prin aceasta se înţelege influenţa asupra lui ξ a unei perturbaţii a coeficienţilor polinomului $p(x)$:

$$p_{\varepsilon}(x) = p(x) + \varepsilon g(x),$$

unde $g(x) \neq 0$ este un polinom arbitrar nenul.

Se poate arăta că dacă ξ este o rădăcină simplă a lui p, atunci pentru pentru valori absolute suficient de mici ale lui ε există o funcţie analitică $\xi(\varepsilon)$, cu $\xi(0) = \xi$, astfel încât $\xi(\varepsilon)$ să fie o rădăcină simplă a polinomului perturbat $p_{\varepsilon}(x)$:

$$p\left(\xi\left(\varepsilon\right)\right) + \varepsilon g\left(\xi(\varepsilon)\right) \equiv 0.$$

De aici, prin derivare în raport cu ε, se obţine pentru $k := \xi'(0)$ ecuaţia

$$kp'\left(\xi(0)\right) + g\left(\xi\left(0\right)\right) = 0,$$

deci

$$k := \frac{-g(\xi)}{p'(\xi)}.$$

Astfel, până la o aproximare de ordinul întâi (adică, eliminând termenii în ε cu puteri mai mari decât 1), se obţine prin dezvoltarea Taylor a lui $\xi(\varepsilon)$ relaţia

$$\xi\left(\varepsilon\right) := \xi - \varepsilon\frac{g\left(\xi\right)}{p'\left(\xi\right)}. \tag{1}$$

În cazul unei rădăcini multiple ξ, de ordin m, se poate arăta că $p(x) + \varepsilon g(x)$ are o rădăcină de forma

$$\xi\left(\varepsilon\right) = \xi + h\left(\varepsilon^{1/m}\right),$$

unde, pentru $|t|$ mic, $h(t)$ este o funcţie analitică cu $h(0) = 0$. Derivând de m ori în raport cu t şi observând că $p(\xi) = p'(\xi) = \cdots = p^{(m-1)}(\xi) = 0$, $p^{(m)}(\xi) \neq 0$, din

$$0 \equiv p_{\varepsilon}\left(\xi\left(\varepsilon\right)\right) = p\left(\xi + h(t)\right) + t^m g\left(\xi + h\left(t\right)\right), \qquad t^m = \varepsilon,$$

pentru $k = h'(0)$ ecuaţia

$$p^{(m)}(\xi)\, k^m + m! g(\xi) = 0,$$

de unde

$$k := \left[-\frac{m! g(\xi)}{p^{(m)}(\xi)}\right]^{1/m}.$$

Din nou, până la o aproximare de ordinul I

$$\xi(\varepsilon) = \xi + \varepsilon^{1/m}\left[-\frac{m! g(\xi)}{p^{(m)}(\xi)}\right]^{1/m}. \tag{2}$$

Pentru $m = 1$ această formulă se reduce la (1) pentru rădăcini simple.

Să presupunem că polinomul $p(x)$ este dat sub formă de listă de coeficienţi

$$p(x) = a_0 x^n + \cdots + a_n.$$

Pentru

$$g_i(x) = a_i x^{n-i}$$

polinomul $p_\varepsilon(x)$ se obţine prin înlocuirea lui a_i în $p(x)$ cu $a_i(1+\varepsilon)$. Formula (2) ne dă următoarea estimare a efectului asupra rădăcinii ξ a unei erori relative ε asupra lui a_i:

$$\xi(\varepsilon) - \xi = \varepsilon^{1/m}\left[-\frac{m! a_i \xi^{n-i}}{p^{(m)}(\xi)}\right]. \tag{3}$$

Astfel, este clar că în cazul rădăcinilor multiple, schimbările în valoarea rădăcinii $\xi(\varepsilon) - \xi$ sunt proporţionale cu $\varepsilon^{1/m}$, $m > 1$, pe când în cazul unei rădăcini simple ele sunt proporţionale doar cu ε: rădăcinile multiple sunt întotdeauna prost condiţionate. Rădăcinile simple pot fi şi ele prost condiţionate. Aceasta se întâmplă dacă factorul lui ε în (3),

$$k(i,\xi) := \left|\frac{a_i \xi^{n-i}}{p'(\xi)}\right|,$$

este mare comparativ cu ξ, ceea ce se poate întâmpla şi în cazul unor polinoame în aparenţă ,,nevătămătoare".

Exemplul 1 (Wilkinson 1959) *Rădăcinile* $\xi_k = k$, $k = 1, 2, \ldots, 20$, *ale polinomului*

$$p(x) = (x-1)(x-2)\ldots(x-20) = \sum_{i=0}^{20} a_i x^{20-i}$$

sunt bine separate. Pentru $\xi_{20} = 20$, *găsim* $p'(20) = 19!$, *şi înlocuind coeficientul* $a_1 = -(1 + 2 + \cdot + 20) = -210$ *cu* $a_1(1+\varepsilon)$ *cauzează o schimbare estimată de*

$$\xi_{20}(\varepsilon) - \xi_{20} = \varepsilon \frac{210 \times 20^{19}}{19!} \approx \varepsilon \cdot 0.9 \times 10^{10}.$$

Cele mai drastice schimbări schimbări sunt cauzate în ξ_{16} *prin perturbarea lui* a_5. *Deoarece* $\xi_{16} = 16$ *şi* $a_5 \approx 10^{-10}$,

$$\xi_{16}(\varepsilon) - \xi_{16} = -\varepsilon a_5 \frac{16^5}{4!15!} \approx \varepsilon \cdot 3.7 \times 10^{14}.$$

Aceasta înseamnă că rădăcinile lui p sunt atât de prost condiţionate încât chiar şi calculul cu o aritmetică de 14 cifre nu va garanta nici o cifră corectă în ξ_{16}.

Exemplul 2 *Prin contrast, rădăcinile polinomului*

$$p(x) = \sum_{i=0}^{20} a_i x^{20-i} := \prod_{j=1}^{20} \left(x - 2^{-j}\right), \qquad \xi_j = 2^{-j},$$

deşi nu sunt bine separate şi se acumulează spre zero, sunt toate bine condiţionate. De exemplu, schimbând a_{20} *în* $a_{20}(1+\varepsilon)$ *cauzează o variaţie a lui* ξ_{20} *care până la o aproximare de ordinul întâi poate fi delimitată sub forma*

$$\left|\frac{\xi_{20}(\varepsilon) - \xi_{20}}{\xi_{20}}\right| = \left|\varepsilon \frac{1}{(2^{-1}-1)(2^{-2}-1)\dots(2^{-19}-1)}\right| \leq 4|\varepsilon|.$$

Mai general, se poate arăta că pentru toate rădăcinile ξ_j *şi toate schimbările* $a_i \to a_i(1+\varepsilon)$

$$\left|\frac{\xi_j(\varepsilon) - \xi_j}{\xi_j}\right| \leq 64|\varepsilon|.$$

Totuşi, rădăcinile sunt bine condiţionate numai în raport cu schimbări relative ale coeficienţilor a_i *şi nu pentru schimbări absolute mici. Dacă înlocuim* $a_{20} = 2^{-210}$ *cu* $\bar{a}_{20} = a_{20} + \Delta a_{20}$, $\Delta a_{20} = 2^{-48}\left(\approx 10^{-14}\right)$ *– acesta poate fi considerată o schimbare absolută mică – atunci polinomul modificat are rădăcinile* $\bar{\xi}_i$ *cu*

$$\bar{\xi}_1 \dots \bar{\xi}_{20} = \overline{a}_{20} = 2^{-120} + 2^{-48} = \left(2^{162}+1\right)\xi_1 \dots \xi_{20}.$$

Cu alte cuvinte, există cel puţin un indice r *cu* $\left|\overline{\xi}_r/\xi_r\right| \geq \left(2^{162}+1\right)^{1/20} > 2^8 = 256$.

Trebuie accentuat că formula (3) se referă doar la senzitivitatea rădăcinilor unui polinom

$$p(x) := \sum_{i=0}^{n} a_i x^{n-i}$$

în reprezentarea uzuală a coeficienţilor. Există şi alte moduri de reprezentare a polinoamelor – de exemplu, ca polinoame caracteristice ale matricelor tridiagonale prin elementele matricei. Efectul unei modificări a parametrilor asupra rădăcinilor într-o astfel de reprezentare pate diferi printr-un ordin de mărime de cel din formula (3) . Condiţionarea rădăcinilor se defineşte întotdeauna cu un tip particular de reprezentare în minte.

Exemplul 3 *Se ştie că pentru orice matrice tridiagonală*

$$J = \begin{bmatrix} \alpha_1 & \beta_2 & & & & \\ \beta_2 & \alpha_2 & \ddots & & & \\ & \ddots & \ddots & \ddots & & \\ & & \ddots & \ddots & \ddots & \\ & & & \ddots & \alpha_{19} & \beta_{20} \\ & & & & \beta_{20} & \alpha_{20} \end{bmatrix}$$

al cărei polinom caracteristic este $p(x) \equiv (x-1)(x-2)\cdots(x-20)$, *schimbări relative mici în* α_i *şi* β_i *cauzează schimbări relative mici ale rădăcinilor* $\xi_j = j$. *În raport cu această reprezentare, toate rădăcinile sunt bine condiţionate, deşi în raport cu reprezentarea uzuală prin coeficienţi sunt foarte prost condiţionate.*

Bibliografie

[1] J. Stoer, R. Bulirsch, Introduction to Numerical Analysis, 2nd edition, Springer, 1992